DRAFT

IS THERE A QUANTUM WORLD?

(Royal Institution Evening Discourse, 28th May 1981)

Mary Hesse

From the time of Descartes to the mid-19th century, most natural philosophers believed that the aim of physics was to describe the real, but hidden, world of fundamental forces and particles, and that these operated like mechanisms. For Descartes, light was refracted like a tennis ball entering a pond, and magnets exerted their force by means of little screw-shaped particles urged along threads within magnetized matter. More than 200 years later William Thompson, Lord Kelvin, said 'I never satisfy myself until I can make a mechanical model of a thing. If I can make a mechanical model I can understand it.'

By the time of Kelvin, however, two of the assumptions of physical mechanism were being questioned. First, it had become incredible to give a realist interpretation of the configurations of mechanisms being used as models in theories of the aether. Maxwell, for instance, described a model in which the magnetic field consisted of fluid vortices along the lines of magnetic force, separated by 'idle particles' like ball-bearings, to take up the rotary motion. These particles in the aether were the electric particles required for Maxwell's controversial displacement current in the aether. He showed how such a mechanism would explain the forces exerted on a current-carrying conductor in a magnetic field. But he was quite explicit that this was only an analogy, not a real description of what the aether is. He says: 'I do not bring it forward as a mode of connexion existing in nature . . . I wish merely to direct the mind of the reader to mechanical phenomena which will assist him in understanding the electrical ones'. Except in one particular

that is, for he goes on 'In speaking of the energy of the field, however, I wish to be understood literally'. The aether, or better the field in space, became a distribution of energy, whatever that is, not a workshop full of mechanical gadgets.

A second assumption was also ready to be dropped: the claim that the whole of nature ought in principle to be reducible to mechanical physics. Amid the current success of subsequent reduction programmes, such as those relating organic processes to biochemistry, and genetics to molecular biology, we are apt to forget that the most ambitious reduction programme in the history of science, namely the reduction of all natural processes to mechanical physics, definitely failed in the 19th century. It failed because physics had to extend its own vocabulary of concepts to incorporate those of electromagnetism, and the relation of mechanics to electromagnetism is not a unified one to this day.

If realism and reductionism can be shown to be untenable in physics, this has important consequences for the philosophical interpretation of science. It means that however good and complete our basic physical theory is, it cannot tell us either what is the fundamental nature of matter, nor what detailed physical laws must be satisfied by more complex organisations of matter, even in biology, let alone in sciences concerning the human mind and human behaviour. For example, philosophy was unnecessarily haunted for centuries by Newtonian determinism; the view that the basic laws of nature determine the future uniquely given its complete mechanical state at any point in time. This doctrine was thought to imply that everything that happens in the natural world, including human behaviour, is uniquely determined by mechanical laws. But this only follows if two conditions are met: first Newtonian mechanics must be accurately true, and

second the whole gamut of natural processes must be totally determined by the entities and laws of physics, that is to say, nature must be reducible to mechanism. We now know that the first condition is not met: modern physics shows that Newtonian mechanism is not accurately true. The doctrine of reductionism has survived, however, in a weakened form. Given that fundamental physics is not mechanistic, we tend to assume that in principle we can find a better theory to replace it, to which all natural processes will be reducible. Whether this is a tenable assumption is a matter of concern to philosophy as well as to physics. I want to argue here that it is not a tenable assumption in quantum theory, and moreover that it is in the end never a sensible assumption for any subsequent theory that might be developed.

Quantum mechanics poses radical difficulties for the programmes of realism and reductionism. What is generally and popularly known about quantum mechanics is not its anti-realist and anti-reductionist implications, but its consequences for determinism. It is well known that the core of the theory is the Uncertainty Principle, according to which simultaneous precise measurements of, for example, the present position and momentum of a quantum particle are impossible. If position is measured relatively accurately, momentum can be determined only relatively inaccurately, and vice versa. It follows immediately that complete dynamical information about the present state of a system can never be obtained, and hence that detailed predictions of its future states are impossible. As far as our measurements are concerned at least, the future is undetermined.

We should notice, however, that a confusion sometimes arises about this consequence of the Uncertainty Principle. It is a confusion that is even found in the standard textbook of philosophy

of science of the 1960s, Ernest Nagel's Structure of Science. Nagel says 'indeterminacy is not exhibited in any experimentally detectable behaviour of macroscopic objects. Indeed the theoretical indeterminism as calculated from quantum mechanics . . . is far smaller than the limits of experimental accuracy'. But Nagel is wrong. It is certainly true that most of the macroscopic objects and systems we experience may be regarded from the point of view of quantum theory as large statistical aggregates of particles, whose behaviour is determinate to well within the limits of experimental accuracy. But there are innumerable kinds of system, both natural and artificial, where single quantum particles may trigger off chains of events in the macro-world, which consequently share and amplify their indeterminism. A homely example is the selection of Premium Bond winners, which may have a significant macroscopic effect on some people's standard of living. Natural examples are disintegrations of small numbers of random atoms, and mutations caused by random radiation, which may have significant macroscopic effects on the course of genetic evolution.

But indeterminism is not the only important consequence of quantum theory. Let us look at the classic two-slit experiment to see what further implications there are for our tottering assumptions of physical realism and physical reductionism. Suppose we have a source of quantum particles, say electrons or photons, which are passed through a diffraction screen with one or two slits, and allowed to fall one by one upon a parallel scintillation screen. With a sufficiently spread out supply of particles, what we observe is a sequence of distinct scintillations on the second screen. Suppose we first let them build up a pattern with the top slit only open, and then with the bottom hole only open. If we are dealing

with ordinary particles, we would expect that when we have both slits open, we would observe a pattern that is just the simple summation, or 'mixture', of the first two patterns. But this is not what we observe - instead the pattern for two slits is what we would expect to observe if classical waves were passing the diffraction screen instead of particles. This is called the 'superposition' pattern. And yet it is not the superposition pattern as a whole that we observe as the 'particles' arrive at the screen one by one, but rather separate scintillations, as if each quantum pattern consists of a wave-packet which 'collapses' instantaneously on to one point when it hits the screen.

What, then, is the 'hidden mechanism' between source and slits and screen? The behaviour of both classical particles and classical waves is inconsistent with observation, so the programmes of classical realism and classical reductionism become untenable. Can we restore these programmes by introducing new types of real hidden entities? This is the programme of so-called 'hidden variables' theories, which seek to find complete descriptions of all quantum states, both observed and unobserved. A complete description of all the hidden variables in the system would make possible both a realistic interpretation of fundamental entities and their laws, and an explanation of observable macro-processes in terms of these entities and laws. This might be possible even though everyone now agrees that such explanations would still be statistical and not deterministic for individuals.

The programme has run into severe difficulties, but before considering these directly, let us see why quantum theory itself as now formulated is incomplete in the sense of not providing us with any such complete hidden variable description. In the two-slit experiment I just outlined, there seems to be an irreducible

discontinuity between what quantum theory describes at the micro-level, and the observable reality at the macro-level. The statistical distribution of particle-waves is given perfectly adequately by Schrödinger's equation when this distribution is observed at the screen. What cannot be explained by Schrödinger's equation is the way the distribution is built up by individual particles one by one. Schrödinger's equation does not tell us that an individual particle has gone through one slit rather than the other. In fact, if we try to catch a particle at one slit or other, we destroy the superposition pattern, and produce instead the mixture of pure particle distributions. What the Schrödinger equation says about an individual particle before it is observed is that it is in a superposition of states, going through slit A and going through slit B, with a probability of $\frac{1}{2}$ each. It is not just that we cannot measure the particle's position, it is also implied by Schrödinger's equation that it has really no position except in this Pickwickian sense of a superposition of probabilities.

This may be all right for unobservable and hypothetical quantum particles whose reality we may doubt anyway, but worse is to follow for the observable macro-world. What follows is vividly illustrated by two striking paradoxes developed early in the history of quantum theory by two of its pioneers, Schrödinger and Einstein.

Schrödinger drew attention to what he called the Cat paradox. Let us go back to the two-slit experiment and suppose that instead of the scintillation screen we put a couple of black boxes behind the slits, in one of which sits a cat. Apart from the slits, which may admit quantum particles one by one from a source, the boxes are sealed from the outside world and from each other. In the top box

where the cat is sitting there is also a lethal device (variously described in the literature as gun or phial of poison) which is guaranteed to be triggered off by entry of a single quantum particle, and if triggered to kill the cat. Now suppose one quantum particle is released, with a probability of $\frac{1}{2}$ of entering slit A, and a probability of $\frac{1}{2}$ of entering slit B. We have just seen that Schrödinger's equation does not say that the unobserved particle has actually entered A and not B, or B and not A. And yet, if the lethal mechanism is working as prescribed, the cat either dies or lives according to whether it did enter A or B, and if we were to do a series of such experiments, we should find approximately half the cats dead and half alive. Before we open up one of the boxes, however, we do not know which of these states a particular cat is in, and yet our deep-rooted assumptions about reality demand that the cat is actually either dead or alive whether we choose to observe it or not. We do not generally subscribe to Bishop Berkeley's theory that the tree ceases to be when no one is about in the quad.

Suppose we also assume, however, in accordance with traditional reducibility, that the macroscopic world too must be explicable by quantum theory, which is after all the most fundamental physical theory around. Then we have to say that both lethal device and cat are in superpositions of states; that is that the cat is both dead and alive with a probability of $\frac{1}{2}$ each. It is tempting to try to evade this consequence by saying that the cat is a highly complex macrosystem consisting of many quantum particles whose behaviour will be governed by quantum statistics rather than by the states of an individual particle. But notice that the experiment has been set up in such a way that we cannot say this: it <u>is</u> the ~~behaviour~~ state of the original single

particle that, according to the Schrödinger equation, determines the consequent state of the cat - this is not a question of large numbers. And Schrödinger's equation says that the cat is in a superposition of states dead-and-alive. What does this mean? The trouble is that it doesn't appear to mean anything, and yet it is the best description we can give not just about what we know but about what there really is according to the Schrödinger equation, there but unobserved.

What has gone wrong? Well, we might try the story about the collapse of the wave-packet. But at what point does the wave-packet collapse? At the quantum particle end of the experiment we undoubtedly have a superposition of states, but at the other end, where 'I' eventually open up the box and observe the state of the cat, there is undoubtedly for me an observation of one state or other. But between these points there is a choice of interpretations: we may say that the collapse takes place (i) at the first macroscopic object (the gun), or (ii) at the first conscious observer (the cat?), or (iii) at myself, who first opens the box. Interpretation (i) has the advantage (if such it can be thought) of brushing reality problems under the carpet and circumventing any challenge to orthodox ways of thinking about the world. But it has the disadvantage of abandoning the reducibility assumption and leaving an ad hoc and quite unexplained hiatus between the micro- and macro-worlds. A 'reality' that is not universalisable is not worth much.

Interpretation (ii) or (iii) might be thought more acceptable in not drawing a line through the physical world, but rather between

the physical world and some 'mind'. It is, of course, generally thought repugnant to physics to ascribe extra-physical effects such as the 'collapse of the wave packet' to the existence of consciousness. The early quantum physicists were quick to point out that it is not so much conscious observation that is involved, but rather the 'measuring apparatus', or any macroscopic system which goes into a definite state on interaction with a micro-system. This response, however, begs the question between the interpretations in favour of interpretation (i), and leads to the so-called 'problem of measurement', which is another version of the problem of non-reducibility. On the other hand, if we take the 'consciousness' interpretation we get into further trouble. Consider an extension of the cat paradox due to Wigner (and hence called 'Wigner's Friend'). Suppose 'you' look into the box and do not tell me the result. Why should I then assume that quantum theory stops short of your brain, so that it is now 9n one of the classical states 'observing cat dead' or 'observing

cat alive'? Consistently with the reducibility assumption I must suppose that your brain too is in a superposition of states, until you tell me the result of your experiment, or I do it myself. Until knowledge is thus shared, according to this view of quantum theory, 'my' world is really different from 'yours' - the common world we can both know exists only when I thus observe it.

Now let us look at Einstein's paradox, called the EPR paradox after Einstein, Podolsky and Rosen who described it in a joint paper in 1935. What follows is a slightly modified version of the thought experiment they presented. It uses the fact that the orthogonal components of the spin of a quantum particle are non-commuting like its position and momentum, and hence cannot be simultaneously measured. Suppose two particles with equal and opposite spin $\pm \frac{1}{2}$ are emitted from a source O at time t_o. At a later time t_A we decide to measure the x-spin of particle A, and find it to be $+\frac{1}{2}$. (We know experimentally that whenever we measure any component spin we shall get $\pm \frac{1}{2}$). Because of the conservation of spin for the particle pair, provided B has not been disturbed in the meantime, we can definitely say that the x-spin of B at t_A and subsequently is $-\frac{1}{2}$, and indeed this prediction has been borne out by experiment.

In classical physics such a situation would not give rise to any problems because it would be assumed that at all times subsequent to t_o A, if undisturbed, actually has x-spin $\frac{1}{2}$ and B has x-spin $-\frac{1}{2}$. and that B's x-spin could have been measured to be $-\frac{1}{2}$ at any time after t_o independent of whether the measurement had been made on A or not. But now consider the following thought experiment. Suppose, after emission at t_o, we had decided what component of spin to measure by tossing a coin or the like, and it had come up not 'measure x-spin'

but 'measure y-spin'. Suppose we had then wheeled up the y-spin measuring apparatus instead of the x-spin apparatus we did actually use, and suppose we had found the y-spin of A to be $+\frac{1}{2}$. From this we could have deduced that the y-spin of B at t_A is $-\frac{1}{2}$. Again classical physics would say, this is because A actually had y-spin $+\frac{1}{2}$ and B had $-\frac{1}{2}$ all the way from O. The correlation between measurements on A and B is a consequence, not of their direct interaction at t_A, but of their previous situation at t_o.

Unfortunately this classical account is inconsistent with quantum theory. Bell and others have shown that there is no way of assigning definite simultaneous values of the x-spin and y-spin components to both particles in the interval t_o to t_A which yields the results of quantum theory. Moreover, actual experiments to reproduce the EPR thought experiment as closely as possible all strongly indicate that quantum theory fits the observations better than any assignment of local hidden variables could do.

These experiments also effectively dispose of another suggestion for evading the EPR paradox. Notice that the paradox depends on what philosophers call 'counterfactual inference'. We have considered what would have happened if we had done other than we actually did, and in conditions where we know that we could not do both what we did and what we might have done, since this would have involved impossible simultaneous measurements. So the alternative to what we did is contrary to fact or 'counterfactual'. Now no paradox arises in regard to what actually happens - the experimental results themselves are not paradoxical, only their theoretical interpretation. So could we not cook up a theoretical interpretation in such a way as to evade paradox? Could we not suppose that when we actually found the x-spin of A to be $+\frac{1}{2}$ and of B to be $-\frac{1}{2}$, these were indeed the spins the particles actually had from t_o? No one can contradict us about this particular case, because the states

of A and B between t_o and t_A have gone for ever and could not have been checked any other way. This is true, but the statistics of similar experiments can contradict us. The experiments that confirm quantum theory rather than hidden variables theory in this context do just that - they are statistical experiments which show that we cannot 'cook up' a hidden variables interpretation in any single case without assuming some sort of conspiracy in the system to satisfy our desires about hidden variables. And the assumption of conspiracies in every single case is inconsistent with both quantum theory and the results of the experiments.

What alternative interpretation of the EPR experiment is available? It appears that the act of measurement on A somehow created the x-spin of $-\frac{1}{2}$ at B when we measured the x-spin at A, and would have created a y-spin of either $\pm\frac{1}{2}$ at B if we had measured the y-spin at A. In other words, particle B somehow 'knew' instantaneously at a distance whether we measured x- or y-spin at A. Now both classical field theory and relativity theory presuppose that causal force can only be propagated by passage of energy from A to B at a velocity not exceeding the velocity of light. The so-called 'non-local' causality of the EPR experiment has to have effects at a distance, instantaneously and without passage of energy. Some intriguing problems arise. First, is this interpretation then inconsistent with the transmission of energy according to relativity? The answer is no, because no energy is transmitted at all. Relativity also asserts that no signal can be transmitted faster than light. Could a couple of spies Arthur and Bernard arrange to communicate from A to B instantaneously by setting up a suitable EPR experiment? The answer again is no, because it is impossible for A and B to arrange both a code-book and the conditions for an interpretable message. If the code instructs Arthur to measure the x-spin to

transmit 'yes' and the y-spin for 'no', then Bernard does not know which test to make to determine the spin for B. If on the other hand the code is $+\frac{1}{2}$ at A for yes and $-\frac{1}{2}$ at A for 'no', then Arthur has no control over which result he gets and therefore cannot use it to send a message. As Michael Redhead puts it, no Bell telephone is made possible by Bell's theorem.

Non-local causality seems to violate many of our most deep-rooted presumptions in physics. Does it not, for example, refute the necessary assumption that we can identify physically isolated systems? Choices as to what measurements to make ~~on~~ at one place may have instantaneous effects elsewhere without any normal means of physical contact between them. These effects may persist over long times and distances and show up in events that are physically isolated from their causes.

In spite of appearances, however, this need not disturb our usual assumptions in physics, for it would be extremely improbable that we would ever notice such classically inexplicable events. Remember that for the instantaneous effects to show up at a distance, it is necessary that particle B has not interacted with any macro-system in the meantime, for if it has its 'memory' of the original correlation with A is lost. It is inconceivable that two particles such as A and B would last long in the real world without interacting with extraneous events and becoming therefore independent systems.

There seems little doubt on both theoretical and experimental grounds that non-local causality exists. Can we, however, still restore the assumptions of realism and reducibility of macro- to micro-physics by means of a hidden variables theory, even if we have to combine it with non-local causality? ~~The answer again is no, and to see why not, let us return briefly to the Cat paradox. There is no difficulty in postulating a set of hidden variables to supplement the Schrödinger~~

The answer seems to be that we cannot restore both realism and reducibility. To see why not, consider Schrodinger's trick of letting a micro-macro interaction take place and then not observing the result, and apply it to the EPR experiment. There is no difficulty in postulating a set of hidden variables equivalent to definite spin states of the two particles in the interval t_o to t_A. The trouble is that if we do that, Bell's theorem tells us that there are no determinist laws connecting these individual states with the states the particles are found to be in when observed. If in the EPR experiment we assume law-like macro-micro reducibility when the particles interact with the apparatus to measure a spin component, then the whole apparatus as well as the particle must be supposed to be in a definite state before any observation is made. When the observation is made, however, this state may change discontinuously and unpredictably. Suppose, for example, that the apparatus at A goes into a state corresponding to a prior x-spin $+\frac{1}{2}$ state of particle A but is not observed. Then we must assume that B has x-spin $-\frac{1}{2}$, and that this will be registered on the apparatus at B when the interaction takes place, say at t_B. But Bell's theorem tells us that it is inconsistent with quantum mechanics to assume that, on every such occasion, the prior x-spin of B was in fact $-\frac{1}{2}$. On some of the occasions when we measure B and observe the result, we shall find its x-spin to be $+\frac{1}{2}$. We shall then have to conclude that the x-spin of A at t_B is $-\frac{1}{2}$, and that the A apparatus is in the corresponding macro-state. This means that a macro-apparatus has to be supposed sometimes to 'jump' from a definite macro-pointer reading $+\frac{1}{2}$ at t_A to $-\frac{1}{2}$ at tB as a result of a non-local event at B. The resulting picture of reality is almost as unacceptable as the notion that superpositions correspond to something real. To avoid this picture we have to reject either the assumption of definite real values, which leads back into the problem of

superpositions, or the assumption of reducibility, If we drop reducibility, then physical laws become functions of the size and nature of a physical system. There is then no need to look to biological organisms for specific irreducible effects of complexity; these effects exist at the heart of physics itself.

I said at the beginning that refutation of the assumptions of realism and/or reducibility would have important consequences for the philosophical interpretation of science. We have seen that if quantum theory is taken to be the most fundamental physical theory, then it is inconsistent with at least one or more of the following assumptions about the world. They are assumptions that are or have been so deeply entrenched that they might be called metaphysical presuppositions:

(1) The future course of the natural world is determined in detail by its present state.

(2) All natural processes are reducible to and explicable by the processes described by the most fundamental physical theory.

(3) The most fundamental physical theory gives a complete description of the real constituents of matter.

(4) Effects can not be transmitted instantaneously accross space without passage of physical energy.

Two final questions arise. First, is quantum theory true? And second, is it reasonable to suppose that we can ever arrive at a fundamental physical theory which will give conclusive answers to questions like those suggested by these four presuppositions?

With regard to the truth of quantum theory, we cannot of course be certain. All that can be said is that any theory which accurately fits the data must be very near to quantum theory, and some have argued that it must be sufficiently near to have at least some of the same consequences with regard to assumptions (1) to (4). But this brings us to the second question. Thomas Kuhn has argued that the historical sequence of fundamental theories or 'paradigms' of nature is never-ending, and that any theory, however well it appears to fit the data at any given time, is liable to run into anomalies requiring revolutionary change of its concepts and laws to produce a new paradigm. The history of physics from Aristotle to quantum theory certainly supports that interpretation. And if it is true, then it seems unwise for philosophers (or even theologians) to worry too much about the transitory implications of current science for metaphysical assumptions such as (1) to (4). One negative thing the history of science does perhaps teach us: we seldom go back exactly to the presuppositions of an earlier science, for example it is unlikely that we shall go back to Newtonian determinism or to classical local causality. But current theories cannot be relied upon to tell us much that is positive that may not be subject to radical reformulation in the future.

I don't know about heaven and earth, but there are certainly more things in physical reality than are dreamed of in our philosophy.